

PROCEDE DE RESTAURATION DE PARTIELS D'UN SIGNAL SONORE

La présente invention se rapporte au domaine des télécommunications et en particulier au domaine du traitement numérique d'un signal sonore et à la représentation harmonique d'un tel signal.

En modélisation harmonique des signaux audionumériques, le signal sonore est représenté par un ensemble d'oscillateurs dont les paramètres (fréquence, amplitude, phase) varient lentement au cours du temps. L'analyse harmonique comprend une analyse temps/fréquence à court terme qui permet de déterminer les valeurs de ces paramètres, suivie d'une extraction de pics puis d'un suivi de partiels.

Le signal à modéliser est découpé en trames de ℓ échantillons (typiquement $\ell = 1024$). Un premier module d'analyse temps/fréquence à court terme (qui effectue typiquement une transformée de Fourier) permet de calculer le spectre du signal à court terme pour chaque trame. Un deuxième module d'extraction de pics permet de ne retenir que les pics les plus pertinents à priori, un critère étant par exemple de ne garder que les pics les plus énergétiques. Un troisième et dernier module cherche à lier les pics entre eux au cours du temps, c'est-à-dire d'une trame à l'autre pour former les partiels. Chaque partiel correspond pendant sa durée de vie à un oscillateur.

Ce type d'analyse et de représentation peut être utilisé en particulier lors d'un codage à réduction de débit, lors d'un codage paramétrique (c'est-à-dire un codage qui traite le signal suivant trois aspects : transitoires, sinusoïdes, bruit), lors de la séparation et l'indexation de sources sonores et lors de la restauration de fichiers sonores.

Il est couramment admis que la synthèse des partiels est de meilleure qualité en utilisant des techniques d'interpolation des phases proposées par Robert J. McAulay et Thomas F. Quatieri dans l'article "Speech Analysis/Synthesis Based on a Sinusoidal Representation", IEEE Transaction on Acoustics, Speech and Signal Processing, PP 744-754, 1986 ou encore proposées par Laurent Girin, Sylvain Marchand, Joseph di Martino, Axel Röbel et Geoffroy Peeters dans l'article "Comparing the order of a Polynomial Phase Model for the Synthesis of Quasi-Harmonic Audio Signals", WASPAA, New Paltz, NY, USA, October 2003. Ces techniques permettent la synthèse d'un partiel d'un pic (A_i, f_i, φ_i) à un pic $(A_{i+1}, f_{i+1}, \varphi_{i+1})$ en calculant toutes les phases intermédiaires à l'aide de polynômes d'ordre 3 ou 5, les fréquences se déduisant

par dérivation. Une interpolation d'ordre 3 est utilisée quand seules sont connues les fréquences et les phases de départ et d'arrivée. Une interpolation d'ordre 5 est utilisée quand sont, en outre, connues les variations d'ordre 2 de la phase (équivalentes aux variations à l'ordre 1 de la fréquence puisque par définition la fréquence est la dérivée de la phase).

La synthèse d'un partiel entre les pics $P_i(A_i, f_i, \varphi_i)$ et $P_{i+1}(A_{i+1}, f_{i+1}, \varphi_{i+1})$ consiste à calculer les valeurs $p(n)$ du partiel entre les trames i et $i+1$:

$$p_i(n) = p(li + n) = A_i(n) \cos(\varphi_i(n)), \quad n = 0, \dots, \ell - 1 \quad (1)$$

A cette fin, il est connu de calculer toutes les phases intermédiaires par l'une des deux méthodes d'interpolation suivantes.

Pour l'interpolation d'ordre 3 selon Mac Aulay et al, la phase est calculée au moyen de l'expression suivante :

$$\varphi_i(n) = \varphi_i + 2\pi f_i nTe + \alpha (nTe)^2 + \beta (nTe)^3 \quad (2)$$

Où Te est la période d'échantillonnage

Les deux inconnues α et β se calculent par résolution d'un système d'équations mettant en jeu $(f_i, \varphi_i, f_{i+1}, \varphi_{i+1})$. Les fréquences se déduisent par dérivation :

$$2\pi f_i(n) = 2\pi f_i + 2\alpha nTe + 3\beta (nTe)^2 \quad (3)$$

Pour l'interpolation d'ordre 5 selon Girin et al, les variations δf_i et δf_{i+1} à l'ordre 1 de la fréquence aux pics P_i et P_{i+1} sont supposées connues. La phase est alors calculée au moyen de l'expression suivante :

$$\varphi_i(n) = \varphi_i + 2\pi f_i nTe + \frac{\delta f_i}{2} (nTe)^2 + \beta (nTe)^3 + \gamma (nTe)^4 + \delta (nTe)^5 \quad (4)$$

Les trois inconnues β, δ, γ se calculent par résolution d'un système d'équations mettant en jeu $(f_i, f_{i+1}, \varphi_i, \varphi_{i+1}, \delta f_i, \delta f_{i+1})$. Les fréquences se déduisent par dérivation :

$$2\pi f_i(n) = 2\pi f_i + \delta f_i nTe + 3\beta (nTe)^2 + 4\gamma (nTe)^3 + 5\delta (nTe)^4 \quad (5)$$

Pour des raisons diverses, il peut arriver que certains partiels existants dans le signal soient absents, corrompus ou discontinus en sortie d'analyse et/ou en entrée de synthèse. Par exemple, ils peuvent être absents en entrée du décodeur dans une

application de diffusion de programmes sonores sur Internet en cas de perte de paquets, ils peuvent être corrompus en cas de perturbations du signal à analyser par un signal parasite (bruit, clic, autre signal, etc), ou ils peuvent être discontinus dans le cas où ils ont une énergie trop faible pour être détectés correctement de façon continue. Il s'avère alors nécessaire de mettre en œuvre des techniques de restauration des pics manquants pour pouvoir recréer un signal synthétisé le plus proche possible du signal original. Ceci nécessite de recréer des pics caractérisés chacun par une amplitude, une fréquence et une phase.

Les techniques précédentes d'interpolation, connues de l'art antérieur, sont utilisées pour synthétiser les parties correspondant aux pics manquants et restaurer les partiels.

Cependant ces techniques d'interpolation connues sont adaptées au court terme, c'est-à-dire pour une période inférieure à 10ms. Pour des durées plus longues, le signal re-synthétisé est souvent éloigné de l'original et des artefacts désagréables peuvent apparaître. En effet, ces techniques assurent une continuité de phase entre les pics existants et les pics restaurés, mais en contre partie elles ne permettent pas de contrôler les fréquences induites données par les équations (3) et (5). Cet effet est d'autant plus marqué que la distance d'interpolation est grande.

Un but de l'invention est de proposer une solution alternative au problème de la restauration de la partie manquante et identifiée comme telle d'un partiel, notamment lorsque la partie manquante correspond à des temps longs (supérieurs à 10 ms) pour lesquels les techniques connues sont peu efficaces.

Aussi, le problème technique à résoudre par l'objet de la présente invention est de proposer un procédé de restauration de parties manquantes des partiels d'un signal sonore, lors d'une analyse harmonique suivant laquelle le signal sonore est découpé en trames temporelles sur lesquelles est appliquée une analyse temps/fréquence qui fournit des spectres successifs à court terme représentés par des trames fréquentielles d'échantillons, l'analyse consistant en outre à extraire des pics spectraux dans les trames fréquentielles et à les lier entre eux au cours du temps pour former des partiels, ce procédé étant une alternative aux solutions connues.

Une solution au problème technique posé consiste, selon la présente invention, en ce que ledit procédé de restauration d'un partiel entre un pic P_i et un pic P_{i+N} dont les fréquences ω et phases ϕ sont connues est tel qu'il comprend les étapes qui consistent :

- à estimer la fréquence $\hat{\omega}$ de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel,
- à calculer la phase déroulée $\hat{\phi}$ de pic en pic, depuis la phase du pic P_i jusqu'à celle du pic P_{i+N} et pour toutes les fréquences $\hat{\omega}$ précédemment estimées,
- à calculer l'erreur de phase $err\phi$ entre la phase déroulée $\hat{\phi}$ et la phase connue au même pic P_{i+N} ,
- à corriger chaque phase déroulée $\hat{\phi}$ d'une valeur fonction de l'erreur de phase $err\phi$.

Un procédé selon l'invention se différencie des méthodes connues en ce qu'il effectue un contrôle plus fin de la fréquence des pics manquants et un calcul après coup des phases correspondantes pour assurer la continuité avec les phases des pics existants. Ainsi, un procédé selon l'invention permet une re-synthèse sans artefacts des signaux correspondants aux morceaux de partiels manquants contrairement aux méthodes connues préalablement décrites.

En outre, de manière avantageuse, un procédé selon l'invention permet une reconstruction du signal plus proche au sens de l'erreur de reconstruction du signal original que celle obtenue par les méthodes connues.

Enfin, un procédé selon l'invention présente avantageusement un algorithme à faible complexité.

L'invention a en outre pour objet un dispositif de synthèse d'un signal sonore pour la mise en œuvre d'un procédé de restauration d'un partiel entre un pic P_i et un pic P_{i+N} . Ce dispositif est par exemple un décodeur audio ou un codeur paramétrique adapté pour la mise en œuvre d'un procédé selon l'invention.

L'invention a en outre pour objet un produit programme d'ordinateur chargeable directement dans la mémoire interne d'un dispositif précédent, ou d'un groupe de dispositifs. Ce produit programme d'ordinateur comprend des portions de code logiciel pour l'exécution des étapes d'un procédé selon l'invention lorsque le programme est exécuté sur le dispositif ou le groupe de dispositifs.

L'invention a en outre pour objet un support utilisable dans un dispositif précédent ou un groupement de dispositifs et sur lequel est enregistré un produit programme d'ordinateur chargeable directement dans la mémoire interne du dispositif ou du groupe de dispositifs, comprenant des portions de code logiciel pour l'exécution des étapes d'un procédé selon l'invention, lorsque le programme est exécuté sur le dispositif ou le groupe de dispositifs.

D'autres caractéristiques et avantages de l'invention apparaîtront lors de la description qui suit faite en regard de figures annexées données à titre d'exemples non limitatifs.

La figure 1 est un organigramme d'un exemple de déroulement d'un procédé selon l'invention.

La figure 2 est un schéma d'un exemple d'utilisation d'un procédé selon l'invention.

Un procédé selon l'invention se déroule de la façon suivante décrite en regard de l'organigramme de la figure 1. Le procédé 1 consiste à restaurer un partiel entre un pic P_i et un pic P_{i+N} dont les fréquences ω et phases φ sont connues.

Dans une première étape 2, le procédé estime la fréquence $\hat{\omega}$ et l'amplitude A de chacun des pics manquants P_{i+1} à P_{i+N-1} . Cette estimation est effectuée par exemple par interpolation ou prédiction linéaire selon des méthodes connues.

Soit un partiel constitué d'une succession de pics $P_i(A_i, \omega_i, \varphi_i)$ liés entre eux, connus à des instants iT et caractérisés par :

A_i , l'amplitude du pic au temps iT

ω_i , la fréquence du pic au temps iT

φ_i , la phase du pic au temps iT , donnée modulo 2π .

L'estimation de la fréquence des pics manquants entre les pics P_i et P_{i+N} est effectuée par exemple au moyen d'une interpolation linéaire entre ω_i et ω_{i+N} , ou au moyen d'une prédiction linéaire sur le passé ou sur le futur, décrite par exemple dans l'article "Enhanced Partial Tracking using linear Prediction", de Mathieu Lagrange, Sylvain Marchand, Martin Raspaud et Jean-Bernard Rault, Proceedings of the Digital Audio Effects (DAFx) Conference, pp141-146, Queen Mary, University of London, UK, September 2003, ou encore au moyen d'une combinaison pondérée sur le passé ou sur le futur.

L'estimation de l'amplitude A des pics manquants est effectuée par exemple au moyen d'une interpolation linéaire entre A_i et A_{i+N} , ou au moyen d'une prédiction linéaire sur le passé ou sur le futur ou encore au moyen d'une combinaison pondérée sur le passé ou sur le futur.

Dans une deuxième étape 3, le procédé calcule la phase déroulée $\hat{\varphi}$ de pic en pic, depuis la phase du pic P_i jusqu'à celle du pic P_{i+N} . Ce calcul est effectué pour chacune des fréquences ω précédemment estimées.

Soient φ_i et ω_i la phase et la fréquence de départ et $\{\hat{\omega}_{i+1}, \dots, \hat{\omega}_{i+N-1}\}$ une estimation des fréquences dans l'intervalle à reconstruire. Pour prolonger le partiel entre le pic P_i et le pic P_{i+N} , le procédé déroule la phase selon l'expression suivante :

$$\hat{\varphi}_{i+n} = \text{mod} \left(\varphi_i + \sum_{j=1}^n \frac{\hat{\omega}_{i+j} + \hat{\omega}_{i+j-1}}{2} T, 2\pi \right), \quad n = 1, \dots, N$$

(6)

Pour ne pas engendrer de discontinuités nuisibles à la qualité de la re-synthèse, il faut obtenir à l'instant $i+N$ une phase reconstruite $\hat{\varphi}_{i+N}$ égale à φ_{i+N} . Les données intervenant dans l'expression (6) précédente étant soit approximées, soit prédites, il est statistiquement impossible d'obtenir cette égalité. Par conséquent, le procédé répartit l'erreur de phase $err\varphi$ calculée à l'instant $i+N$ entre tous les pics manquants et préalablement reconstruits P_{i+1} à P_{i+N-1} au moyen des étapes suivantes.

Dans une troisième étape 4, le procédé calcule l'erreur de phase $err\varphi$ entre la phase déroulée $\hat{\varphi}_{i+N}$ et la phase connue φ_{i+N} au même pic P_{i+N} . Ce calcul peut être effectué selon le système d'équations suivant :

$$err\varphi = \varphi_{i+N} - \hat{\varphi}_{i+N} + 2\pi \quad \text{si} \quad |\varphi_{i+N} - \hat{\varphi}_{i+N} + 2\pi| < |\varphi_{i+N} - \hat{\varphi}_{i+N}|$$

(7)

$$err\varphi = \varphi_{i+N} - \hat{\varphi}_{i+N} - 2\pi \quad \text{si} \quad |\varphi_{i+N} - \hat{\varphi}_{i+N} - 2\pi| < |\varphi_{i+N} - \hat{\varphi}_{i+N}|$$

(8)

$$err\varphi = \varphi_{i+N} - \hat{\varphi}_{i+N} \quad \text{sinon.}$$

(9)

Dans une quatrième étape 5, le procédé corrige chaque phase déroulée $\hat{\varphi}_{i+n}$ d'une valeur fonction de l'erreur de phase $err\varphi$. Typiquement l'erreur de phase calculée à l'instant $i+N$ est répartie de manière uniforme sur chacune des phases déroulées selon l'expression suivante :

$$\hat{\varphi}_{corrigée_{i+n}} = \text{mod} \left(\hat{\varphi}_{i+n} + err\varphi \frac{n}{N}, 2\pi \right) \quad n = 1, \dots, N-1$$

(10)

La répartition peut ne pas être uniforme et suivre une loi non linéaire par exemple.

L'exemple d'utilisation illustré par la figure 2, consiste à restaurer les partiels au moyen d'un procédé 1 selon l'invention lors d'une analyse harmonique d'un signal sonore par exemple lors d'un codage paramétrique. Le signal sonore $s(n)$ est représenté par un ensemble d'oscillateurs dont les paramètres (fréquence, amplitude) varient lentement au cours du temps. De manière classique, l'analyse harmonique comprend une analyse 6 temps/fréquence à court terme qui permet de déterminer les valeurs de ces paramètres, suivie d'une extraction 7 de pics puis d'un suivi 8 de partiels. Une détection 9 de trous dans les partiels précède la mise en œuvre d'un procédé 1 de restauration des partiels selon l'invention. Les pics reconstruits $P_{i+n}(\hat{A}_{i+n}, \hat{\omega}_{i+n}, \hat{\phi}_{i+n})$ lors de la mise en œuvre du procédé 1, sont ensuite traités comme des pics issus de l'analyse harmonique et la synthèse 10 additive du signal correspondant au partiel restauré à partir de ces pics reconstruits peut se faire, par exemple, par une des méthodes connues d'interpolation des phases (ordre 3 ou ordre 5).

REVENDICATIONS

1. Procédé (1) de restauration de partiels d'un signal sonore, lors d'une analyse harmonique suivant laquelle le signal sonore est découpé en trames temporelles sur lesquelles est appliquée une analyse temps/fréquence qui fournit des spectres successifs à court terme représentés par des trames fréquentielles d'échantillons, l'analyse consistant en outre à extraire des pics spectraux dans les trames fréquentielles et à les lier entre eux au cours du temps pour former des partiels, le procédé de restauration d'un partiel entre un pic P_i et un pic P_{i+N} dont les fréquences et phases sont connues est caractérisé en ce qu'il comprend les étapes qui consistent :

- à estimer (2) la fréquence $\hat{\omega}$ de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel,
- à calculer (3) la phase déroulée $\hat{\phi}$ de pic en pic, depuis la phase du pic P_i jusqu'à celle du pic P_{i+N} et ce, pour toutes les fréquences $\hat{\omega}$ précédemment estimées,
- à calculer (4) l'erreur de phase $err\phi$ entre la phase déroulée $\hat{\phi}$ et la phase connue au même pic P_{i+N} ,
- à corriger (5) chaque phase déroulée $\hat{\phi}$ d'une valeur fonction de l'erreur de phase $err\phi$.

2. Procédé (1) de restauration de partiels d'un signal sonore selon la revendication 1, dans lequel la phase déroulée $\hat{\phi}$ est calculée par la formule :

$$\hat{\phi}_{i+n} = \text{mod} \left(\phi_i + \sum_{j=1}^n \frac{\hat{\omega}_{i+j} + \hat{\omega}_{i+j-1}}{2} T, 2\pi \right), \quad n = 1, \dots, N$$

avec ϕ_i et $\hat{\omega}_i = \omega_i$ la phase et la fréquence du pic P_i , ϕ_{i+N} et $\hat{\omega}_{i+N} = \omega_{i+N}$ la phase et la fréquence du pic P_{i+N} .

3. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 et 2, dans lequel l'estimation de la fréquence $\hat{\omega}$ des pics manquants P_{i+1} à P_{i+N-1} est effectuée au moyen d'une interpolation linéaire entre les fréquences des pics connus P_i et P_{i+N} .

4. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 et 2, dans lequel l'estimation de la fréquence $\hat{\omega}$ des pics manquants P_{i+1} à P_{i+N-1} est effectuée au moyen d'une prédiction linéaire sur le passé.
5. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 et 2, dans lequel l'estimation de la fréquence $\hat{\omega}$ des pics manquants P_{i+1} à P_{i+N-1} est effectuée au moyen d'une prédiction linéaire sur le futur.
6. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 et 2, dans lequel l'estimation de la fréquence $\hat{\omega}$ des pics manquants P_{i+1} à P_{i+N-1} est effectuée au moyen d'une combinaison pondérée d'une prédiction linéaire sur le passé et d'une prédiction linéaire sur le futur.
7. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications précédentes, qui comprend en outre l'étape qui consiste :
 - à estimer l'amplitude de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel au moyen d'une interpolation linéaire entre les amplitudes A des pics connus P_i et P_{i+N} .
8. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 à 6, qui comprend en outre l'étape qui consiste :
 - à estimer l'amplitude de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel au moyen d'une prédiction linéaire sur le passé.
9. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 à 6, qui comprend en outre l'étape qui consiste :
 - à estimer l'amplitude de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel au moyen d'une prédiction linéaire sur le futur.
10. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications 1 à 6, qui comprend en outre l'étape qui consiste :

- à estimer l'amplitude de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel au moyen d'une prédiction linéaire sur le passé et d'une prédiction linéaire sur le futur.

11. Procédé (1) de restauration de partiels d'un signal sonore selon l'une des revendications précédentes, dans lequel la correction de phase consiste à répartir uniformément l'erreur de phase $err\varphi$ calculée à l'instant $i+N$ entre tous les pics manquants P_{i+1} à P_{i+N-1} du partiel.

12. Procédé (1) de restauration de partiels d'un signal sonore selon la revendication précédente, dans lequel la phase corrigée est déterminée par l'équation :

$$\hat{\varphi}_{corrigée_{i+n}} = \text{mod} \left(\hat{\varphi}_{i+n} + err\varphi \frac{n}{N}, 2\pi \right) \quad n=1, \dots, N-1$$

13. Procédé (1) de restauration de partiels d'un signal sonore selon la revendication précédente, dans lequel l'erreur de phase $err\varphi$ est déterminée par le système d'équations :

$$\begin{aligned} err\varphi &= \varphi_{i+N} - \hat{\varphi}_{i+N} + 2\pi \quad \text{si } |\varphi_{i+N} - \hat{\varphi}_{i+N} + 2\pi| < |\varphi_{i+N} - \hat{\varphi}_{i+N}| \\ err\varphi &= \varphi_{i+N} - \hat{\varphi}_{i+N} - 2\pi \quad \text{si } |\varphi_{i+N} - \hat{\varphi}_{i+N} - 2\pi| < |\varphi_{i+N} - \hat{\varphi}_{i+N}| \\ err\varphi &= \varphi_{i+N} - \hat{\varphi}_{i+N} \quad \text{sinon.} \end{aligned}$$

14. Dispositif de synthèse d'un signal sonore pour la mise en œuvre d'un procédé selon l'une des revendications précédentes, caractérisé en ce qu'il comprend :

- des moyens d'estimation de la fréquence $\hat{\omega}$ de chacun des pics manquants P_{i+1} à P_{i+N-1} de ce partiel,
- des moyens de calcul de la phase déroulée $\hat{\varphi}$ de pic en pic, depuis la phase du pic P_i jusqu'à celle du pic P_{i+N} et ce, pour toutes les fréquences $\hat{\omega}$ précédemment estimées,
- des moyens de calcul de l'erreur de phase $err\varphi$ entre la phase déroulée $\hat{\varphi}$ et la phase connue au même pic P_{i+N} ,
- des moyens de correction de chaque phase déroulée $\hat{\varphi}$ d'une valeur fonction de l'erreur de phase $err\varphi$.

15. Un produit programme d'ordinateur chargeable directement dans la mémoire interne d'un dispositif ou d'un groupe de dispositifs selon la revendication

précédente, comprenant des portions de code logiciel pour l'exécution des étapes d'un procédé (1) selon l'une des revendications 1 à 13, lorsque le programme est exécuté sur le dispositif ou le groupe de dispositifs.

16. Support utilisable dans un dispositif ou un groupement de dispositifs selon la revendication 14 et sur lequel est enregistré un produit programme d'ordinateur chargeable directement dans la mémoire interne du dispositif ou du groupe de dispositifs, comprenant des portions de code logiciel pour l'exécution des étapes d'un procédé (1) selon l'une des revendications 1 à 13, lorsque le programme est exécuté sur le dispositif ou le groupe de dispositifs.